

RAW SEQUENCE LISTING

**The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.**

Application Serial Number: 10/ 009, 873
Source: IFW16
Date Processed by STIC: 04/07/2006

ENTERED



IFW16

RAW SEQUENCE LISTING

DATE: 04/07/2006

PATENT APPLICATION: US/10/009,873

TIME: 08:31:48

Input Set : A:\Sequence Listing.TXT

Output Set: N:\CRF4\04072006\J009873.raw

3 <110> APPLICANT: RANGANATHAN, ANAND
 5 <120> TITLE OF INVENTION: DNA MANIPULATION METHODS AND APPLICATIONS FOR SYNTHETIC
 6 ENZYMES
 8 <130> FILE REFERENCE: SHW-009US
 10 <140> CURRENT APPLICATION NUMBER: 10/009,873
 11 <141> CURRENT FILING DATE: 2002-11-1
 13 <150> PRIOR APPLICATION NUMBER: PCT/GB00/02286
 14 <151> PRIOR FILING DATE: 2000-06-12
 16 <150> PRIOR APPLICATION NUMBER: GB 9913694.7
 17 <151> PRIOR FILING DATE: 1999-06-11
 19 <160> NUMBER OF SEQ ID NOS: 30
 21 <170> SOFTWARE: PatentIn Ver. 3.3
 23 <210> SEQ ID NO: 1
 24 <211> LENGTH: 12
 25 <212> TYPE: DNA
 26 <213> ORGANISM: Artificial Sequence
 28 <220> FEATURE:
 29 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
 30 oligonucleotide
 32 <400> SEQUENCE: 1
 33 aaatctagac cc 12
 36 <210> SEQ ID NO: 2
 37 <211> LENGTH: 12
 38 <212> TYPE: DNA
 39 <213> ORGANISM: Artificial Sequence
 41 <220> FEATURE:
 42 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
 43 oligonucleotide
 45 <400> SEQUENCE: 2
 46 agatctagac cc 12
 49 <210> SEQ ID NO: 3
 50 <211> LENGTH: 374
 51 <212> TYPE: PRT
 52 <213> ORGANISM: Streptomyces lividans
 54 <400> SEQUENCE: 3
 55 Met Ala Gly Thr Asp Arg Glu Lys Ala Leu Asp Ala Ala Leu Ala Gln
 56 1 5 10 15
 58 Ile Glu Arg Gln Phe Gly Lys Gly Ala Val Met Arg Met Gly Asp Arg
 59 20 25 30
 61 Thr Asn Glu Pro Ile Glu Val Ile Pro Thr Gly Ser Thr Ala Leu Asp
 62 35 40 45
 64 Val Ala Leu Gly Val Gly Gly Ile Pro Arg Gly Arg Val Val Glu Val
 65 50 55 60

CP9-6)

RAW SEQUENCE LISTING

DATE: 04/07/2006

PATENT APPLICATION: US/10/009,873

TIME: 08:31:48

Input Set : A:\Sequence Listing.TXT

Output Set: N:\CRF4\04072006\J009873.raw

```

67 Tyr Gly Pro Glu Ser Ser Gly Lys Thr Thr Leu Thr Leu His Ala Val
68 65 70 75 80
70 Ala Asn Ala Gln Lys Ala Gly Gly Gln Val Ala Phe Val Asp Ala Glu
71 85 90 95
73 His Ala Leu Asp Pro Glu Tyr Ala Lys Lys Leu Gly Val Asp Ile Asp
74 100 105 110
76 Asn Leu Ile Leu Ser Gln Pro Asp Asn Gly Glu Gln Ala Leu Glu Ile
77 115 120 125
79 Val Asp Met Leu Val Arg Ser Gly Ala Leu Asp Leu Ile Val Ile Asp
80 130 135 140
82 Ser Val Ala Ala Leu Val Pro Arg Ala Glu Ile Glu Gly Glu Met Gly
83 145 150 155 160
85 Asp Ser His Val Gly Leu Gln Ala Arg Leu Met Ser Gln Ala Leu Arg
86 165 170 175
88 Lys Ile Thr Ser Ala Leu Asn Gln Ser Lys Thr Thr Ala Ile Phe Ile
89 180 185 190
91 Asn Gln Leu Arg Glu Lys Ile Gly Val Met Phe Gly Ser Pro Glu Thr
92 195 200 205
94 Thr Thr Gly Gly Arg Ala Leu Lys Phe Tyr Ala Ser Val Arg Leu Asp
95 210 215 220
97 Ile Arg Arg Ile Glu Thr Leu Lys Asp Gly Thr Asp Ala Val Gly Asn
98 225 230 235 240
100 Arg Thr Arg Val Lys Val Val Lys Asn Lys Val Ala Pro Pro Phe Lys
101 245 250 255
103 Gln Ala Glu Phe Asp Ile Leu Tyr Gly Gln Gly Ile Ser Arg Glu Gly
104 260 265 270
106 Gly Leu Ile Asp Met Gly Val Glu Asn Gly Phe Val Arg Lys Ala Gly
107 275 280 285
109 Ala Trp Tyr Thr Tyr Glu Gly Asp Gln Leu Gly Gln Gly Lys Glu Asn
110 290 295 300
112 Ala Arg Asn Phe Leu Lys Asp Asn Pro Asp Leu Ala Asn Glu Ile Glu
113 305 310 315 320
115 Lys Lys Ile Lys Gln Lys Leu Gly Val Gly Val His Pro Glu Glu Ser
116 325 330 335
118 Ala Thr Glu Pro Gly Ala Asp Ala Ala Ser Ala Ala Pro Ala Asp Ala
119 340 345 350
121 Ala Pro Ala Val Pro Ala Pro Thr Thr Ala Lys Ala Thr Lys Ser Lys
122 355 360 365
124 Ala Ala Ala Ala Lys Ser
125 370
128 <210> SEQ ID NO: 4
129 <211> LENGTH: 372
130 <212> TYPE: PRT
131 <213> ORGANISM: Streptomyces ambofaciens
133 <400> SEQUENCE: 4
134 Met Ala Gly Thr Asp Arg Glu Lys Ala Leu Asp Ala Ala Leu Ala Gln
135 1 5 10 15
137 Ile Glu Arg Gln Phe Gly Lys Gly Ala Val Met Arg Met Gly Asp Arg
138 20 25 30

```

RAW SEQUENCE LISTING

DATE: 04/07/2006

PATENT APPLICATION: US/10/009,873

TIME: 08:31:48

Input Set : A:\Sequence Listing.TXT

Output Set: N:\CRF4\04072006\J009873.raw

```

140 Ser Lys Glu Pro Ile Glu Val Ile Pro Thr Gly Ser Thr Ala Leu Asp
141      35      40      45
143 Val Ala Leu Gly Val Gly Gly Leu Pro Arg Gly Arg Val Ile Glu Val
144      50      55      60
146 Tyr Gly Pro Glu Ser Ser Gly Lys Thr Thr Leu Thr Leu His Ala Val
147 65      70      75      80
149 Ala Asn Ala Gln Lys Ala Gly Gly Gln Val Ala Phe Val Asp Ala Glu
150      85      90      95
152 His Ala Leu Asp Pro Glu Tyr Ala Gln Lys Leu Gly Val Asp Ile Asp
153      100      105      110
155 Asn Leu Ile Leu Ser Gln Pro Asp Asn Gly Glu Gln Ala Leu Glu Ile
156      115      120      125
158 Val Asp Met Leu Val Arg Ser Gly Ala Leu Asp Leu Ile Val Ile Asp
159      130      135      140
161 Ser Val Ala Ala Leu Val Pro Arg Ala Glu Ile Glu Gly Glu Met Gly
162 145      150      155      160
164 Asp Ser His Val Gly Leu Gln Ala Arg Leu Met Ser Gln Ala Leu Arg
165      165      170      175
167 Lys Ile Thr Ser Ala Leu Asn Gln Ser Lys Thr Thr Ala Ile Phe Ile
168      180      185      190
170 Asn Gln Leu Arg Glu Lys Ile Gly Val Met Phe Gly Ser Pro Glu Thr
171      195      200      205
173 Thr Thr Gly Gly Arg Ala Leu Lys Phe Tyr Ala Ser Val Arg Leu Asp
174      210      215      220
176 Ile Arg Arg Ile Glu Thr Leu Lys Asp Gly Thr Asp Ala Val Gly Asn
177 225      230      235      240
179 Arg Thr Arg Val Lys Val Val Lys Asn Lys Val Ala Pro Pro Phe Lys
180      245      250      255
182 Gln Ala Glu Phe Asp Ile Leu Tyr Gly Gln Gly Ile Ser Arg Glu Gly
183      260      265      270
185 Gly Leu Ile Asp Met Gly Val Glu His Gly Phe Val Arg Lys Ala Gly
186      275      280      285
188 Ala Trp Tyr Thr Tyr Glu Gly Asp Gln Leu Gly Gln Gly Lys Glu Asn
189      290      295      300
191 Ala Arg Asn Phe Leu Lys Asp Asn Pro Asp Leu Ala Asn Glu Ile Glu
192 305      310      315      320
194 Lys Lys Ile Lys Glu Lys Leu Gly Val Gly Val Arg Pro Glu Glu Pro
195      325      330      335
197 Thr Ala Thr Glu Ser Gly Pro Asp Ala Ala Thr Ala Glu Ser Ala Pro
198      340      345      350
200 Ala Val Pro Ala Pro Ala Thr Ala Lys Val Thr Lys Ala Lys Ala Ala
201      355      360      365
203 Ala Ala Lys Ser
204      370
207 <210> SEQ ID NO: 5
208 <211> LENGTH: 1125
209 <212> TYPE: DNA
210 <213> ORGANISM: Streptomyces lividans
212 <400> SEQUENCE: 5

```

RAW SEQUENCE LISTING

DATE: 04/07/2006

PATENT APPLICATION: US/10/009,873

TIME: 08:31:48

Input Set : A:\Sequence Listing.TXT

Output Set: N:\CRF4\04072006\J009873.raw

```

213 atggcaggaa cgcaccgcga gaaggccctg gacgccgcgc tcgcacagat tgaacggcaa 60
214 ttccggcaagg gcgcgggtcat gcgcattgggt gaccggacca acgagcccat cgaggtcatc 120
215 ccgaccgggt ctaccgcgct cgacgtggcc ctccggctcg gaggcattcc gcgtggccgt 180
216 gtcgtggagg tctacggccc cgagtcctcg ggcaagacga ccctgaccct gcacgcggtg 240
217 gcgaacgcgc agaaggccgg cggccagggtc gcgttcgtgg acgccgagca cgccctcgac 300
218 cccgagtacg cgaagaagct cgggtgtcgac atcgacaacc tgatcctgtc ccagccggac 360
219 aacggtgagc aggccctgga gatcgtggac atgctggtcc gctccggcgc cctcgacctc 420
220 atcgtcatcg actccgtcgc cgcgctcgtc ccgcgcgcgg agatcgaggg cgagatgggc 480
221 gacagccacg tcggtctgca ggcccggctg atgagccagg ccctgcggaa gatcaccagc 540
222 gcgctcaacc agtccaagac caccgcgac ttcatcaacc agctccgcga gaagatcggc 600
223 gtgatgttcg gtcctccgga gaccacgacc ggtggccggg cactgaagtt ctacgcctcg 660
224 gtgcgactcg acatccggcg tatcgagacg ctgaaggacg gcaccgacgc ggtcggcaac 720
225 cgcacccgcg tcaaggtggt caagaacaag gtcgcgcgcg ccttcaagca ggccgagttc 780
226 gacatcctct acggccaggg catcagccgc gagggcgggtc tgatcgacat gggcgtggag 840
227 aacggcttcg tccgcaaggc cggcgccctg tacacgtacg agggcgacca gctcggtcag 900
228 ggcaaggaga acgcgcgcaa cttcctgaag gacaaccccg acctggcaa cgagatcgag 960
229 aagaagatca agcagaagct gggcgctcggc gtgcaccccg aggagtcggc caccgagccc 1020
230 ggcgcggacg ccgcctccgc cgcgccggcc gacgccgcac cggcggtgcc cgcacccacg 1080
231 accgccaagg ccaccaagtc caaggccgcg gcagccaaga gctga 1125
234 <210> SEQ ID NO: 6
235 <211> LENGTH: 1119
236 <212> TYPE: DNA
237 <213> ORGANISM: Streptomyces ambofaciens
239 <400> SEQUENCE: 6
240 atggcaggaa cgcaccgcga gaaggctctt gacgccgcac tcgcacagat tgaacggcag 60
241 ttccggcaagg gcgcgggtcat gcgcattgggt gaccgggtcga aggagcccat cgaggtcatc 120
242 ccgaccgggt cgaccgcgct cgacgtggcc ctccggctcg gcggcctgcc gcgcggccgc 180
243 gtcgtggagg tctacgggtc ggagtcctcc ggttaagacga ccctgaccct gcacgcggtg 240
244 gcgaacgcgc agaaggccgg cggccagggtg gcgttcgtgg acgccgagca cgccctcgac 300
245 cccgagtacg cccagaagct cggcgctcgac atcgacaacc tgatcctgtc ccagccggac 360
246 aacggtgagc aggccctgga gatcgtggac atgctggtcc gctccggcgc cctcgacctc 420
247 atcgtcatcg actccgtcgc cgcgctcgtc ccgcgcgcgg agatcgaggg cgagatgggt 480
248 gacagccacg tcggtctcca ggcccggctg atgagccagg cgctccggaa gatcaccagc 540
249 gcgctcaacc agtccaagac caccgcgac ttcatcaacc agctccgcga gaagatcggc 600
250 gtcatgttcg gtcctccgga gaccacgacc ggtggccggg cgctcaagtt ctacgcctcg 660
251 gtgcgactcg acatccgacg catcgagacg ctcaaggacg gcaccgacgc ggtcggcaac 720
252 cgcacgcgcg tcaaggtcgt caagaacaag gtcgcgcgcg ccttcaagca ggccgagttc 780
253 gacatcctct acggccaggg catcagccgc gagggcggcc tgatcgacat gggcgtggag 840
254 cacggcttcg tccgcaaggc cggcgccctg tacacgtacg agggcgacca gctcggccag 900
255 ggcaaggaga acgcgcgcaa cttcctgaag gacaaccccg acctcgcaa cgagatcgag 960
256 aagaagatca aggagaagct gggcgctcggc gtccgtcccg aggagccgac ggccaccgag 1020
257 tccggaccgg acgccgcgac ggccgaatcc gcaccggcgg tgcccgcgcc cgcgaccgcc 1080
258 aaggtcacca aggccaaggc cgcggcagcc aagagctga 1119
261 <210> SEQ ID NO: 7
262 <211> LENGTH: 14
263 <212> TYPE: DNA
264 <213> ORGANISM: Artificial Sequence
266 <220> FEATURE:
267 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic

```

RAW SEQUENCE LISTING

DATE: 04/07/2006

PATENT APPLICATION: US/10/009,873

TIME: 08:31:48

Input Set : A:\Sequence Listing.TXT

Output Set: N:\CRF4\04072006\J009873.raw

```

268      oligonucleotide
270 <220> FEATURE:
271 <221> NAME/KEY: modified_base
272 <222> LOCATION: (6)..(14)
273 <223> OTHER INFORMATION: a, c, g, t, unknown or other
275 <400> SEQUENCE: 7
W--> 276 ggatgnnnnn nnnn                                14
279 <210> SEQ ID NO: 8
280 <211> LENGTH: 18
281 <212> TYPE: DNA
282 <213> ORGANISM: Artificial Sequence
284 <220> FEATURE:
285 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
286      oligonucleotide
288 <220> FEATURE:
289 <221> NAME/KEY: modified_base
290 <222> LOCATION: (1)..(13)
291 <223> OTHER INFORMATION: a, c, g, t, unknown or other
293 <400> SEQUENCE: 8
W--> 294 nnnnnnnnnn nnncatcc                                18
297 <210> SEQ ID NO: 9
298 <211> LENGTH: 35
299 <212> TYPE: DNA
300 <213> ORGANISM: Artificial Sequence
302 <220> FEATURE:
303 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
304      oligonucleotide
306 <400> SEQUENCE: 9
307 ggcatatggc ggacctgtca aagctctccg acagt                35
310 <210> SEQ ID NO: 10
311 <211> LENGTH: 36
312 <212> TYPE: DNA
313 <213> ORGANISM: Artificial Sequence
315 <220> FEATURE:
316 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
317      oligonucleotide
319 <400> SEQUENCE: 10
320 ggtctagatc ccagccgcgg tcggtcggca gtcccc                36
323 <210> SEQ ID NO: 11
324 <211> LENGTH: 53
325 <212> TYPE: DNA
326 <213> ORGANISM: Artificial Sequence
328 <220> FEATURE:
329 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
330      oligonucleotide
332 <400> SEQUENCE: 11
333 ggtctagact cgctgttcca ccccgacccc acgcgctcgg gcaccgcgca cca        53
336 <210> SEQ ID NO: 12
337 <211> LENGTH: 48

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/009,873

DATE: 04/07/2006
TIME: 08:31:49

Input Set : A:\Sequence Listing.TXT
Output Set: N:\CRF4\04072006\J009873.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:7; N Pos. 6,7,8,9,10,11,12,13,14

Seq#:8; N Pos. 1,2,3,4,5,6,7,8,9,10,11,12,13

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/009,873

DATE: 04/07/2006

TIME: 08:31:49

Input Set : A:\Sequence Listing.TXT

Output Set: N:\CRF4\04072006\J009873.raw

L:11 M:256 W: Invalid Numeric Header Field, Wrong Current FILING DATE:YYYY-MM-DD

L:276 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7 after pos.:0

L:294 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:0